



IFW16

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/09/874,390A

TIME: 12:30:03

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\09132004\I874390A.raw

3 <110> APPLICANT: CLAUSEN, Henrik

5 <120> TITLE OF INVENTION: UDP-N-ACETYLGLUCOSAMINE: GALACTOSE-B1, 3-N-

ACETYLGLACTOSAMINE-a-R / N-

6 ACETYL GLUCOSAMINE-B1,3-N-ACETYLGLACTOSAMINE-a-R (GlcNAc to GalNAc) B1,6-N-

7 ACETYLGLUCOSAMINYLTRANSFERASE, C2/4GnT

9 <130> FILE REFERENCE: 4305/0J425

11 <140> CURRENT APPLICATION NUMBER: US 09/874,390A

12 <141> CURRENT FILING DATE: 2001-06-04

14 <150> PRIOR APPLICATION NUMBER: DK PA 1988 01605

15 <151> PRIOR FILING DATE: 1998-12-04

17 <160> NUMBER OF SEQ ID NOS: 12

19 <170> SOFTWARE: PatentIn version 3.1

21 <210> SEQ ID NO: 1

22 <211> LENGTH: 2319

23 <212> TYPE: DNA

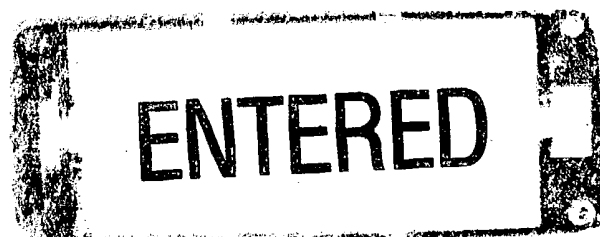
24 <213> ORGANISM: Homo sapiens

26 <220> FEATURE:

27 <221> NAME/KEY: CDS

28 <222> LOCATION: (496)..(1809)

29 <223> OTHER INFORMATION:



W--> 31 <400> 1

```

32 attaactggg ttttctatt tatctatcct ctgcgattac ttctctgagt cagagcctct      60
34 tctctctaag tcacgggaac tgcccttgct acttgtagacc tgccctttac tcagcagttt      120
36 ttgttctggg aagccttggg attctgctaa tacctatcac tgtaggtgct gaagggaac      180
38 agatgaagaa catgacctca aggagcttcc tgtcaatgag aagaccaagc tgacgcctgg      240
40 caaagatatt aaagaggagc ctgaaactgt tccttggaaca tcttatgaat gtcagaaaat      300
42 accttttggg ggggttagaag atcaggggac atggttggtc acatttgctg ccacggaaca      360
44 ccgccagtct tcacttggaac acagaatcac gccttggtgaa gagatcatcc ctaagcagga      420
46 gagaagctac taaaggattg tgcctctctc caccttccct gtgctcggtc tccacctgtc      480
48 tcccattctg tgacg atg gtt caa tgg aag aga ctc tgc cag ctg cat tac      531
49 Met Val Gln Trp Lys Arg Leu Cys Gln Leu His Tyr
50 1 5 10
52 ttg tgg gct ctg ggc tgc tat atg ctg ctg gcc act gtg gct ctg aaa      579
53 Leu Trp Ala Leu Gly Cys Tyr Met Leu Leu Ala Thr Val Ala Leu Lys
54 15 20 25
56 ctt tct ttc agg ttg aag tgt gac tct gac cac ttg ggt ctg gag tcc      627
57 Leu Ser Phe Arg Leu Lys Cys Asp Ser Asp His Leu Gly Leu Glu Ser
58 30 35 40
60 agg gaa tct caa agc cag tac tgt agg aat atc ttg tat aat ttc ctg      675
61 Arg Glu Ser Gln Ser Gln Tyr Cys Arg Asn Ile Leu Tyr Asn Phe Leu
62 45 50 55 60
64 aaa ctt cca gca aag agg tct atc aac tgt tca ggg gtc acc cga ggg      723
65 Lys Leu Pro Ala Lys Arg Ser Ile Asn Cys Ser Gly Val Thr Arg Gly
66 65 70 75

```

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/09/874,390A

TIME: 12:30:03

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\09132004\I874390A.raw

68	gac	caa	gag	gca	gtg	ctt	cag	gct	att	ctg	aat	aac	ctg	gag	gtc	aag	771
69	Asp	Gln	Glu	Ala	Val	Leu	Gln	Ala	Ile	Leu	Asn	Asn	Leu	Glu	Val	Lys	
70				80						85				90			
72	aag	aag	cga	gag	cct	ttc	aca	gac	acc	cac	tac	ctc	tcc	ctc	acc	aga	819
73	Lys	Lys	Arg	Glu	Pro	Phe	Thr	Asp	Thr	His	Tyr	Leu	Ser	Leu	Thr	Arg	
74			95					100					105				
76	gac	tgt	gag	cac	ttc	aag	gct	gaa	agg	aag	ttc	ata	cag	ttc	cca	ctg	867
77	Asp	Cys	Glu	His	Phe	Lys	Ala	Glu	Arg	Lys	Phe	Ile	Gln	Phe	Pro	Leu	
78		110					115					120					
80	agc	aaa	gaa	gag	gtg	gag	ttc	cct	att	gca	tac	tct	atg	gtg	att	cat	915
81	Ser	Lys	Glu	Glu	Val	Glu	Phe	Pro	Ile	Ala	Tyr	Ser	Met	Val	Ile	His	
82	125					130					135				140		
84	gag	aag	att	gaa	aac	ttt	gaa	agg	cta	ctg	cga	gct	gtg	tat	gcc	cct	963
85	Glu	Lys	Ile	Glu	Asn	Phe	Glu	Arg	Leu	Leu	Arg	Ala	Val	Tyr	Ala	Pro	
86				145						150				155			
88	cag	aac	ata	tac	tgt	gtc	cat	gtg	gat	gag	aag	tcc	cca	gaa	act	ttc	1011
89	Gln	Asn	Ile	Tyr	Cys	Val	His	Val	Asp	Glu	Lys	Ser	Pro	Glu	Thr	Phe	
90			160					165					170				
92	aaa	gag	gag	gtc	aaa	gca	att	att	tct	tgc	ttc	cca	aat	gtc	ttc	ata	1059
93	Lys	Glu	Ala	Val	Lys	Ala	Ile	Ile	Ser	Cys	Phe	Pro	Asn	Val	Phe	Ile	
94			175					180					185				
96	gcc	agt	aag	ctg	gtt	cgg	gtg	gtt	tat	gcc	tcc	tgg	tcc	agg	gtg	caa	1107
97	Ala	Ser	Lys	Leu	Val	Arg	Val	Val	Tyr	Ala	Ser	Trp	Ser	Arg	Val	Gln	
98		190					195					200					
100	gct	gac	ctc	aac	tgc	atg	gaa	gac	ttg	ctc	cag	agc	tca	gtg	ccg	tgg	1155
101	Ala	Asp	Leu	Asn	Cys	Met	Glu	Asp	Leu	Leu	Gln	Ser	Ser	Val	Pro	Trp	
102	205				210						215				220		
104	aaa	tac	ttc	ctg	aat	aca	tgt	ggg	acg	gac	ttt	cct	ata	aag	agc	aat	1203
105	Lys	Tyr	Phe	Leu	Asn	Thr	Cys	Gly	Thr	Asp	Phe	Pro	Ile	Lys	Ser	Asn	
106				225						230				235			
108	gca	gag	atg	gtc	cag	gct	ctc	aag	atg	ttg	aat	ggg	agg	aat	agc	atg	1251
109	Ala	Glu	Met	Val	Gln	Ala	Leu	Lys	Met	Leu	Asn	Gly	Arg	Asn	Ser	Met	
110			240					245					250				
112	gag	tca	gag	gta	cct	cct	aag	cac	aaa	gaa	acc	cgc	tgg	aaa	tat	cac	1299
113	Glu	Ser	Glu	Val	Pro	Pro	Lys	His	Lys	Glu	Thr	Arg	Trp	Lys	Tyr	His	
114			255					260					265				
116	ttt	gag	gta	gtg	aga	gac	aca	tta	cac	cta	acc	aac	aag	aag	aag	gat	1347
117	Phe	Glu	Val	Val	Arg	Asp	Thr	Leu	His	Leu	Thr	Asn	Lys	Lys	Lys	Asp	
118		270					275					280					
120	cct	ccc	cct	tat	aat	tta	act	atg	ttt	aca	ggg	aat	gag	tac	att	gtg	1395
121	Pro	Pro	Pro	Tyr	Asn	Leu	Thr	Met	Phe	Thr	Gly	Asn	Ala	Tyr	Ile	Val	
122	285					290					295				300		
124	gct	tcc	cga	gat	ttc	gtc	caa	cat	gtt	ttg	aag	aac	cct	aaa	tcc	caa	1443
125	Ala	Ser	Arg	Asp	Phe	Val	Gln	His	Val	Leu	Lys	Asn	Pro	Lys	Ser	Gln	
126			305							310				315			
128	caa	ctg	att	gaa	tgg	gta	aaa	gac	act	tat	agc	cca	gat	gaa	cac	ctc	1491
129	Gln	Leu	Ile	Glu	Trp	Val	Lys	Asp	Thr	Tyr	Ser	Pro	Asp	Glu	His	Leu	
130			320					325					330				
132	tgg	gcc	acc	ctt	cag	cgt	gca	cgg	tgg	atg	cct	ggc	tct	gtt	ccc	aac	1539

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/09/874,390A

TIME: 12:30:03

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\09132004\I874390A.raw

```

133 Trp Ala Thr Leu Gln Arg Ala Arg Trp Met Pro Gly Ser Val Pro Asn
134          335                      340                      345
136 cac ccc aag tac gac atc tca gac atg act tct att gcc agg ctg gtc      1587
137 His Pro Lys Tyr Asp Ile Ser Asp Met Thr Ser Ile Ala Arg Leu Val
138          350                      355                      360
140 aag tgg cag ggt cat gag gga gac atc gat aag ggt gct cct tat gct      1635
141 Lys Trp Gln Gly His Glu Gly Asp Ile Asp Lys Gly Ala Pro Tyr Ala
142 365                      370                      375                      380
144 ccc tgc tct gga atc cac cag cgg gct atc tgc gtt tat ggg gct ggg      1683
145 Pro Cys Ser Gly Ile His Gln Arg Ala Ile Cys Val Tyr Gly Ala Gly
146          385                      390                      395
148 gac ttg aat tgg atg ctt caa aac cat cac ctg ttg gcc aac aag ttt      1731
149 Asp Leu Asn Trp Met Leu Gln Asn His His Leu Leu Ala Asn Lys Phe
150          400                      405                      410
152 gac cca aag gta gat gat aat gct ctt cag tgc tta gaa gaa tac cta      1779
153 Asp Pro Lys Val Asp Asp Asn Ala Leu Gln Cys Leu Glu Glu Tyr Leu
154          415                      420                      425
156 cgt tat aag gcc atc tat ggg act gaa ctt tgagacacac tatgagagcg      1829
157 Arg Tyr Lys Ala Ile Tyr Gly Thr Glu Leu
158          430                      435
160 ttgctacctg tggggcaaga gcatgtacaa acatgctcag aacttgctgg gacagtgtgg      1889
162 gtgggagacc agggctttgc aattcgtggc atccttttagg ataagagggc tgctattaga      1949
164 ttgtgggtaa gtagatcttt tgccttgcaa attgctgcct gggatgaatgc tgcttgttct      2009
166 ctcaccctta accctagtag ttcctccact aactttctca ctaagtgaga atgagaactg      2069
168 ctgtgatagg gagagtgaag gagggatatg tggtagagca cttgatttca gttgaatgcc      2129
170 tgctggtagc ttttccattc tgtggagctg ccgttcctaa taattccagg tttggtagcg      2189
172 tggaggagaa ctttgatgga aagagaacct tcccttctgt actgttaact taaaaataaa      2249
174 tagctctga ttcaaagtat tacctctact ttttgcttag tatgccagaa ataataataa      2309
176 tctaaacaga      2319
179 <210> SEQ ID NO: 2
180 <211> LENGTH: 438
181 <212> TYPE: PRT
182 <213> ORGANISM: Homo sapiens
184 <400> SEQUENCE: 2
186 Met Val Gln Trp Lys Arg Leu Cys Gln Leu His Tyr Leu Trp Ala Leu
187 1          5          10          15
190 Gly Cys Tyr Met Leu Leu Ala Thr Val Ala Leu Lys Leu Ser Phe Arg
191          20          25          30
194 Leu Lys Cys Asp Ser Asp His Leu Gly Leu Glu Ser Arg Glu Ser Gln
195          35          40          45
198 Ser Gln Tyr Cys Arg Asn Ile Leu Tyr Asn Phe Leu Lys Leu Pro Ala
199          50          55          60
202 Lys Arg Ser Ile Asn Cys Ser Gly Val Thr Arg Gly Asp Gln Glu Ala
203 65          70          75          80
206 Val Leu Gln Ala Ile Leu Asn Asn Leu Glu Val Lys Lys Lys Arg Glu
207          85          90          95
210 Pro Phe Thr Asp Thr His Tyr Leu Ser Leu Thr Arg Asp Cys Glu His
211          100         105         110
214 Phe Lys Ala Glu Arg Lys Phe Ile Gln Phe Pro Leu Ser Lys Glu Glu

```

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/09/874,390A

TIME: 12:30:03

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\09132004\I874390A.raw

```

215          115          120          125
218 Val Glu Phe Pro Ile Ala Tyr Ser Met Val Ile His Glu Lys Ile Glu
219          130          135          140
222 Asn Phe Glu Arg Leu Leu Arg Ala Val Tyr Ala Pro Gln Asn Ile Tyr
223 145          150          155          160
226 Cys Val His Val Asp Glu Lys Ser Pro Glu Thr Phe Lys Glu Ala Val
227          165          170          175
230 Lys Ala Ile Ile Ser Cys Phe Pro Asn Val Phe Ile Ala Ser Lys Leu
231          180          185          190
234 Val Arg Val Val Tyr Ala Ser Trp Ser Arg Val Gln Ala Asp Leu Asn
235          195          200          205
238 Cys Met Glu Asp Leu Leu Gln Ser Ser Val Pro Trp Lys Tyr Phe Leu
239          210          215          220
242 Asn Thr Cys Gly Thr Asp Phe Pro Ile Lys Ser Asn Ala Glu Met Val
243 225          230          235          240
246 Gln Ala Leu Lys Met Leu Asn Gly Arg Asn Ser Met Glu Ser Glu Val
247          245          250          255
250 Pro Pro Lys His Lys Glu Thr Arg Trp Lys Tyr His Phe Glu Val Val
251          260          265          270
254 Arg Asp Thr Leu His Leu Thr Asn Lys Lys Lys Asp Pro Pro Pro Tyr
255          275          280          285
258 Asn Leu Thr Met Phe Thr Gly Asn Ala Tyr Ile Val Ala Ser Arg Asp
259          290          295          300
262 Phe Val Gln His Val Leu Lys Asn Pro Lys Ser Gln Gln Leu Ile Glu
263 305          310          315          320
266 Trp Val Lys Asp Thr Tyr Ser Pro Asp Glu His Leu Trp Ala Thr Leu
267          325          330          335
270 Gln Arg Ala Arg Trp Met Pro Gly Ser Val Pro Asn His Pro Lys Tyr
271          340          345          350
274 Asp Ile Ser Asp Met Thr Ser Ile Ala Arg Leu Val Lys Trp Gln Gly
275          355          360          365
278 His Glu Gly Asp Ile Asp Lys Gly Ala Pro Tyr Ala Pro Cys Ser Gly
279          370          375          380
282 Ile His Gln Arg Ala Ile Cys Val Tyr Gly Ala Gly Asp Leu Asn Trp
283 385          390          395          400
286 Met Leu Gln Asn His His Leu Leu Ala Asn Lys Phe Asp Pro Lys Val
287          405          410          415
290 Asp Asp Asn Ala Leu Gln Cys Leu Glu Glu Tyr Leu Arg Tyr Lys Ala
291          420          425          430
294 Ile Tyr Gly Thr Glu Leu
295          435
298 <210> SEQ ID NO: 3
299 <211> LENGTH: 21
300 <212> TYPE: DNA
301 <213> ORGANISM: Artificial Sequence
303 <220> FEATURE:
304 <223> OTHER INFORMATION: PCR primer TSHC37
306 <400> SEQUENCE: 3
307 ggaagttcat acagttccca c

```

21

RAW SEQUENCE LISTING

DATE: 09/13/2004

PATENT APPLICATION: US/09/874,390A

TIME: 12:30:03

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\09132004\I874390A.raw

```

310 <210> SEQ ID NO: 4
311 <211> LENGTH: 21
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: PCR primer TSHC28
318 <400> SEQUENCE: 4
319 cctcccatc aacatcttga g                                21
322 <210> SEQ ID NO: 5
323 <211> LENGTH: 31
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: PCR primer TSHC45
330 <400> SEQUENCE: 5
331 agcgaattca gctcaaagtt cagtcccata g                                31
334 <210> SEQ ID NO: 6
335 <211> LENGTH: 21
336 <212> TYPE: DNA
337 <213> ORGANISM: Artificial Sequence
339 <220> FEATURE:
340 <223> OTHER INFORMATION: PCR primer TSHC48
342 <400> SEQUENCE: 6
343 gtgggaactg tatgaacttc c                                21
346 <210> SEQ ID NO: 7
347 <211> LENGTH: 30
348 <212> TYPE: DNA
349 <213> ORGANISM: Artificial Sequence
351 <220> FEATURE:
352 <223> OTHER INFORMATION: PCR primer TSHC54
354 <400> SEQUENCE: 7
355 gcagaattca tggttcaatg gaagagactc                                30
358 <210> SEQ ID NO: 8
359 <211> LENGTH: 26
360 <212> TYPE: DNA
361 <213> ORGANISM: Artificial Sequence
363 <220> FEATURE:
364 <223> OTHER INFORMATION: PCR primer TSHC55
366 <400> SEQUENCE: 8
367 cgagaattca ggttgaagtg tgactc                                26
370 <210> SEQ ID NO: 9
371 <211> LENGTH: 21
372 <212> TYPE: DNA
373 <213> ORGANISM: Artificial Sequence
375 <220> FEATURE:
376 <223> OTHER INFORMATION: PCR primer TSHC78
378 <400> SEQUENCE: 9
379 gctcggcttc caccgtgtctc c                                21
382 <210> SEQ ID NO: 10

```

VERIFICATION SUMMARY

DATE: 09/13/2004

PATENT APPLICATION: US/09/874,390A

TIME: 12:30:04

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\09132004\I874390A.raw

L:31 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:29